scientific reports

Check for updates

OPEN

Identification of novel *ABCB4* variants and genotype-phenotype correlation in progressive familial intrahepatic cholestasis type 3

Senyan Wang[1,2,5], Qi Liu[2,5], Xiaoyan Sun[3], Wenjuan Wei[1], Leilei Ding[4]✉ & Xiaofang Zhao[1]✉

Progressive familial intrahepatic cholestasis type 3 (PFIC3) is a severe hepatic disorder characterized by cholestasis. Elucidating the genotype-phenotype correlations and expanding the mutational spectrum of the *ABCB4* gene are crucial for enhancing diagnostic accuracy and therapeutic strategies. Clinical and genetic data from 2 original PFIC3 patients from our institution, along with 118 additional cases identified through a comprehensive literature review, were integrated for a comprehensive analysis. The study included statistical analysis of clinical information, genetic analysis, multi-species sequence alignment, protein structure modeling, and pathogenicity assessment. Machine learning techniques were applied to identify genotype-phenotype relationships. We identified three novel *ABCB4* mutations: two missense mutations (c.904G > T and c.2493G > C) and one splicing mutation (c.1230 + 1G > A). Homozygous mutations were associated with significantly earlier disease onset compared to compound heterozygous mutations ($p < 0.0001$). Missense mutations were predominant (76.9%), with Exon 7 being the most frequently affected region. A random forest model indicated that Exon 10 had the highest feature importance score (9.9%). Liver transplantation remains the most effective treatment modality for PFIC3. This investigation broadens the known mutation spectrum of the *ABCB4* gene and identifies key variant sites associated with clinical manifestations. These insights lay a foundation for early diagnosis, optimal treatment selection, and further research into PFIC3.

Keywords PFIC3, *ABCB4* gene, Liver transplantation; machine learning, Mutation spectrum

Abbreviations

PFIC3	Progressive familial intrahepatic cholestasis type 3
UDCA	Ursodeoxycholic acid
MDR3	Multidrug Resistance Protein 3
PEBD	Partial external biliary drainage

Background

Progressive familial intrahepatic cholestasis type 3 (PFIC3) is an uncommon and potentially lethal autosomal recessive disorder characterized by early onset and persistent cholestasis. Jaundice, pruritus, abdominal pain, steatorrhea, anorexia and growth retardation usually occur from infancy to early childhood, eventually progressing to cirrhosis and liver failure[1–3]. The incidence of PFIC3 is estimated to be about 1 in 50,000 to 100,000 live births, varying by race and region[4]. The disease is more prevalent in families with a history of consanguineous marriage[1].

PFIC3 is caused by variants in the *ABCB4*gene, which encodes Multidrug Resistance Protein 3 (MDR3), a biliary ATP-binding cassette (ABC) protein responsible for the transport of phosphatidylcholine (PC) from

[1]Translational Medicine Centre, The First Affiliated Hospital of Zhengzhou University, Zhengzhou University, Zhengzhou 450000, China. [2]Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China. [3]Department of Oncology, Henan Cancer Hospital, Zhengzhou University Affiliated Cancer Hospital, Zhengzhou, China. [4]Department of Obstetrics and Gynecology, Peking Union Medical College, National Clinical Research Center for Obstetric & Gynecologic Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China. [5]Senyan Wang and Qi Liu contributed equally to this work. ✉email: leileiding0520@163.com; zhaoxiaofang2099@163.com

hepatocytes into bile[5–7]. Variants in the *ABCB4*gene result in aberrant expression or dysfunction of the MDR3 protein, thereby leading to impaired phospholipid secretion, resulting in the inability of bile salts to be inactivated by phospholipids. This pathological process causes damage to the epithelial cells of the bile ducts and biliary tree, ultimately leading to cholestasis, cholangitis, biliary constriction, and biliary cirrhosis[6,8,9].

Definitive diagnosis of PFIC3 often requires genetic testing. Ursodeoxycholic acid, recognized as the first-line therapeutic intervention for all PFIC subtypes, demonstrating particular efficacy in ameliorating the symptoms of milder PFIC3 presentations[10]. Novel gene therapy approaches and small molecule drugs have not yet been validated in clinical care[11,12]. It is important to note that while pharmacological interventions may ameliorate symptoms, they are not curative, liver transplantation often represents the sole definitive treatment option[13]. Different types of variants (e.g. missense, nonsense, splice-site variants, etc.) and different variant sites have been found to have different effects on the function of the MDR3 protein, and can independently influence clinical manifestations and disease progression[14]. Therefore, the reporting of new variant sites and the study of the relationship between variant sites and the clinical phenotypes provide assistance in understanding the pathogenesis of PFIC3.

In this study, we reported two PFIC3 patients from Chinese families and alongside a retrospective analysis of 118 previously reported PFIC3 cases. We described in detail the clinical features, genetic profiles, gene expression patterns, and diagnostic process of the patients. Our variant analyses identified three novel and unreported variants: c.904G > T/p.G302C (P1), c.2493G > C/p.R831S (P2-parental origin), and c.1230 + 1G > A (splice site) (P2-parental origin), which extended the variant spectrum of the *ABCB4* gene (see Table 1). In addition, we investigated the relationship between the variant location of the *ABCB4* gene and the age of clinical onset using a random forest classification machine learning approach, which revealed key variant sites. These discoveries contribute significantly to the genetic understanding of PFIC3, and provide novel perspectives for clinical diagnosis and personalized treatment.

Methods

Patients

Taking into account the patients' medical history, signs, clinical findings and genetic testing, we excluded a series of diseases that could potentially induce cholestasis, including viral hepatitis, alcoholic liver disease, nonalcoholic fatty liver disease, other hereditary cholestatic liver diseases, biliary atresia, Alagille syndrome, α1-antitrypsin deficiency, cystic fibrosis, sclerosing cholangitis, and intra- and extrahepatic obstruction. We confirmed the diagnosis of two patients with PFIC3 and collected information on their clinical family pedigrees and drew genealogical maps.The study was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (Approval No. 2022-KY-0826-001), and all family members or legal guardians of patients under 18 years of age signed a written informed consent, informed consent was obtained from all participants. Confirming that all experiments were performed in accordance with relevant guidelines and regulations.The data in this study were anonymized or treated confidentially and did not affect the diagnosis or treatment of the patients.

Clinical evaluations

The retrospective analysis encompassed the patients' clinical data, comprising gender, age, age at initial visit, and the presence of jaundice, hepatomegaly, splenomegaly, cirrhosis, portal hypertension, and gallstones, which were analyzed retrospectively. Biochemical test data, including Alanine Transaminase (ALT), Aspartate Aminotransferase (AST), serum alkaline phosphatase (ALP), glutamyltransferase (GGT), total bilirubin

Patient	Chromosome coordinates	Exon	Gene subregion	Nucleotide change/ Mutant protein	Type of variation	Genetic model	Source of variation	Variation classificat-ion	PolyPhen*/ SIFT*/ Provean*/ MutationTaster*
P1*	Chr7: 87,037,496	Exon_25	CDS24	c.3136 C > T/ p.R1046X	nonsense	AR*	Mother	Patho-genic	-/ -/ D(-14.07/ D
	Chr7: 87,076,451	Exon_9	CDS8	c.904G > T/ p.G302C	missense	AR*	Father	VUS*	D(1.0)/ D(0.00)/ D(-7.58)/D
P2*	Chr7: 87,046,817	Exon_21	CDS20	c.2493G>/ p.R831S	missense	AR*	Mother	VUS*	D(1.0)/ D(0.001)/D(-5.06)/D
	Chr7: 87,072,978	Exon_12	CDS11	c.1230 + 1G > A/ splicing	splicing	AR*	Father	Likely Patho-genic	-/ -/ -/ -

Table 1. Genetic variation information of both patients. *Genetic model: AR (Autosomal Recessive). Variation classification: VUS (Variants of Uncertain Significance). PolyPhen: (http://genetics.bwh.harvard.edu/pph2/.) D: Probably damaging (pp2_hdiv ≥ 0.957), P: Possibly damaging (0.453 ≤ pp2_hdiv ≤ 0.956), B: Benign (pp2_hdiv ≤ 0.452). SIFT/Provean: (http://provean.jcvi.org/protein_batch_submit.php?species=human.) D: Damaging (sift ≤ 0.05), T: Tolerated (sift > 0.05) / D: Deleterious (-14 < provean<-2.5), T: Tolerated (-2.5 < provean < 14). MutationTaster: (http://www.mutationtaster.org/.) A: Disease causing_automatic, D: Disease_causing, N: Polymorphism, P: Polymorphism_automatic. Patient: P1 (Patient 1), P2 (Patient 2).

(TBiL), direct bilirubin (DBiL), and total bile acids (TBA), were collected. All patients underwent abdominal ultrasonography (US) at our center.

Variant analysis and protein function prediction

We extracted genomic DNA from the provided samples and prepared DNA libraries using the hybridization capture method after fragmentation, ligation of junctions, amplification and purification. Then, the high-throughput sequencing platform was used to analyze the exonic regions of 20,099 genes in the human whole exome and the intronic regions 20 bp next to them. The sequencing data were aligned with the human genome reference sequence hg19 (GRCh37), and the coverage and sequencing quality of the target regions were assessed. Pathogenicity of variants was assessed according to the 2015 edition of the American College of Medical Genetics and Genomics (ACMG) guidelines[15]. Blood DNA samples were collected from pre-certified individuals and their family members, including parents and siblings, for genetic variant analysis via sequencing, thereby substantiating the genetic diagnoses. We predicted the deleteriousness of the variants using the SIFT[16], PolyPhen-2[17], Mutation Taster[18] and PROVEAN[19] programs. Predictions were made using the Human Splicing Finder tool, which can assess the likelihood of splice site variants disrupting normal splicing (http://www.umd.be/HSF3/HSF.shtml)[20].

Amino acid sequence conservation analysis

To elucidate the authenticity of the impairment of MDR3 activity caused by missense variants, we conducted an analysis of the evolutionary conservation of the mutated amino acids. This was accomplished through a comparative alignment of amino acid sequences from *ABCB4*homologs, obtained from the Ensembl Genome Browser. This analysis was performed by MEGA XI (Molecular Evolutionary Genetics Analysis) software[21]. The 10 primate homologous proteins include Homo sapiens (P21439), Pan troglodytes (H2R3D1), Mus musculus (P21440), Rattus norvegicus (Q08201), Sus scrofa (I3LV60), Loxodonta africana (G3TGB0), Felis catus (A0A2I2UPK2), Ficedula albicollis (U3JZD2), Canis lupus familiaris (A0A8I3N6T3), and Mustela putorius furo (M3Z3M8).

Protein structure construction and variant analysis

The 3D structures of proteins were downloaded from the Protein Data Bank (PDB) via PyMOL (version 2.4.1) and visualized. Amino acid variants were introduced at specified positions using PyMOL's variant tool, and energy minimization was performed to optimize the structure. The functional implications of these variants were evaluated by meticulously comparing the structural alterations, both pre- and post-variant, with a particular focus on critical functional domains and active sites.

Literature review

The keywords "ABCB4" and "PFIC3" were searched in the PubMed and CNKI databases to collect and analyze genotypic and phenotypic data from patients with a clear molecular genetic diagnosis and detailed clinical information. The final analysis included 118 patients with PFIC3. Data were sourced from 38 publications, consisting of 1 review article, 1 multicenter study, 22 case reports, and 14 research articles. These diverse sources provided a comprehensive dataset that includes both individual case details and broader, multi-institutional patient samples, ensuring a thorough analysis of the genotype-phenotype correlation in PFIC3.

Statistical analysis methods

The data were collected using WPS Office, analyzed and plotted using the GraphPad Prism (version 8.0.1) and IBM SPSS Statistics 26. A random forest classification model was constructed from the training set data, and feature importance was calculated. Normally distributed data were expressed as mean ± standard deviation (mean ± SD) and compared using t-tests; categorical variables were expressed as percentages. All statistical tests were performed with $p < 0.05$ as the criterion of significance.

Results

Case overview and clinical characterization

The case study at our center included two patients: one female and one male. Patient 1 was a female who presented with jaundice as the first symptom at 16 months of age. Patient 2 was a male, initially presented at the age of 13 years without any notable clinical signs(see Table 1); however, his first clinical manifestation was the presence of abnormally elevated liver function tests. Both patients demonstrated pronounced hepatic dysfunction, evidenced by elevated levels of biochemical markers such as AST, ALT, TBiL, DBiL, TBA, ALP, and GGT, which are indicative of cholestasis (see Table 2). After a detailed investigation, we excluded a range of common etiologies, including biliary atresia, congenital biliary dilatation, viral hepatitis, and other metabolic disorders, including citrulline deficiency, benign recurrent intrahepatic cholestasis, and bile acid synthesis defects, were excluded. Abdominal ultrasonography of patient 1 disclosed hepatomegaly and splenomegaly. The abdominal ultrasonographic examination of Patient 2 indicated cirrhosis and portal hypertension, with characteristic features including splenomegaly and ascites. To ascertain the definitive diagnosis, both patients underwent comprehensive genetic testing (see Table 1).

Analysis of familial inheritance and clinical significance of complex heterozygous variants of *ABCB4* gene

Both patients exhibited the presence of complex heterozygous variants of the *ABCB4* gene. Four distinct variants of *ABCB4* were identified, three of which were previously uncharacterized, and one of which was previously

	Gender/Age (years)	Jaundice	Portal hyperten-sion	Time		TBil* μmol/L 0–25	DBil* μmol/L 0–10	TBA* μmo/ L 0–20	ALT* U/L 0–40	AST* U/L 0–40	ALP* U/L 0-281	GGT* U/L 0–58	Ceruloplasmin mg/dL 15–45
				State	Day								
P1*	Female/ 1.5	Yellow staining of the sclera	Splenomegaly	Before- LT*	Last test	47.9	26.9	235.6	82	238	247	126	47.77
				Post-orthotopic- LT*	1d	67.6	43.9	/	497	540	97	44	/
					2d	27.5	18.8	/	678	502	83	38	/
					3d	17.7	10.3	/	448	229	78	140	/
					4d	18.5	12.9	/	300	105	75	43	/
					5d	20.6	14.6	37.4	212	59	103	49	17.17
					8d	12.5	6.7	14.2	88	55	96	26	13.83
					12d	4.4	2.7	5.9	62	68	80	41	18.59
					14d	4	2.5	4.9	43	69	96	41	18.18
					16d	3.9	1.4	7.2	32	47	103	38	25.01
					18d	5	2.6	5.2	19	32	119	28	31.41
					20d	3.7	1.3	5.6	15	29	128	19	/
P2*	Male/ 23	Yellow staining of the sclera	Splenomegaly, varicose veins of the esophagus	Before- LT	Last test	628	465.3	260.6	40	80	154	53	25.16
				Post-orthotopic- LT	1d	423.1	323.2	/	182	384	103	41	/
					2d	195.7	160.5	10	161	169	90	38	25.27
					3d	172.6	145.2	8.2	116	59	87	52	25.78
					4d	148	124.6	32	93	38	99	113	27.38
					5d	122.3	98.3	15	77	41	87	96	19.63
					6d	130.7	102.1	29.9	69	37	103	110	20.46
					7d	148.1	122.9	154.4	60	36	144	159	22.93
					8d	156.7	130	179.8	50	29	186	173	22.74
					9d	114.7	90.1	50.3	44	25	209	168	23.82
					10d	79.3	66.4	21.8	73	57	169	135	20.44
					11d	74.3	59.1	10.7	68	35	147	106	18.42
					12d	65.8	55.5	32.4	54	19	135	95	18.23
					155d	13.5	5.4	16.3	17	17	297	21	\
					180d	10.3	5.2	12	10	14	363	20	\
					216d	19.2	7.7	16.1	14	17	3113	17	\

Table 2. Liver function indicators of patients before and after liver transplantation. *TBil: Total serum bilirubin.DBil: Direct bilirubin.TBA: Total serum bile acid.ALT: Alanine transaminase. AST: Aspartate aminotransferase. ALP: Serum alkaline phosphatase. GGT: Serum gamma glutamyltransferase. P1: Patient 1.P2: Patient 2. LT: Liver transplantation.

known. Two variants were classified as pathogenic (P) or likely pathogenic (LP) according to the ACMG guidelines, and the other two variants were defined as Variants of Uncertain Significance (VUS). The results of the trio whole-exome sequencing (Trio-WES) indicated that all of the observed variants were inherited from the parents. The inheritance patterns for each patient are detailed in Table 1.

Patient 1 carries a complex heterozygous variant of the *ABCB4* gene (c.3136 C > T, c.904G > T) (Fig. 1A). The nonsense variant c.3136 C > T (p.R1046X) leads to a change from C to T at base 3136 in the coding region, converting the 1046th amino acid from arginine to a stop codon, resulting in premature termination of the *ABCB4* protein translation (Fig. 1B). This variant has been reported in the literature and has been assessed as pathogenic by ACMG (P). The missense variant c.904G > T (p.G302C) results in a change from G to T at the 904th base in the coding region, changing the 302nd amino acid from glycine to cysteine (Fig. 1B). This variant has not been previously documented in the scientific literature and was evaluated by ACMG as VUS. Sequencing confirmed that the mother was heterozygous for the c.3136 C > T (p.R1046X) variant, while the father was heterozygous for the c.904G > T (p.G302C) variant (Fig. 2A, B).

Patient 2 also carried a complex heterozygous variant of *ABCB4* gene (c.1230 + 1G > A, c.2493G > C) (Fig. 1A). The splice variant c.1230 + 1G > A is not reported in the HGMDpro database and has been classified as potentially pathogenic by ACMG (LP). The missense variant c.2493G > C (p.R831S) results in a change from G to C at base 2493 within the coding region, which consequently alters amino acid 831 from arginine to serine (Fig. 1B). This variant has not been previously documented in the scientific literature and has been evaluated by the ACMG as VUS. Trio-WES results showed that the patient's mother carries the c.2493G > C (p.R831S) variant, the father carries the c.1230 + 1G > A splice site variant, and the fraternal twin brother also carries the c.1230 + 1G > A splice site variant, while the younger brother does not carry these gene variants (Fig. 2C, D).

Computer simulation analysis of three new *ABCB4* gene variants evaluated

Multiple sequence alignment analysis was performed to assess the evolutionary conservatism across different species. Our results showed that the three variants in *ABCB4*, c.904G > T (p.G302C), c.1230 + 1G > A (affecting splicing), and c.2493G > C (p.R831S), all affected strictly conserved domains in orthologous MDR3 proteins, including those in distantly related species (Fig. 1C). These results suggest that variants at these three amino acid sites may alter the protein structure and lead to functional deficiency of the MDR3 protein.

The bioinformatics software Splice AI was used to predict the effect of the c.1230 + 1G > A splicing variant on protein splicing, suggesting that the original splice recognition site was disrupted, leading to the use of an alternative splicing donor site (Fig. 1D).

The other two novel missense variants were also predicted to be deleterious by four prediction tools, including Polyphen-2, SIFT, Provean, and Mutation Taster (Table 1). We further examined the tertiary structure of the proteins with two missense variants defined as VUS. As shown in Fig. 1E, the spatial structure of the peptide chain around the variant amino acid was significantly altered. This change in the spatial structure of the protein may lead to the abnormal transport function of the MDR3 protein.

Analysis of the relationship between the variation pattern and site of *ABCB4* gene and the clinical phenotype of PFIC3 disease

To further clarify the relationship between the type and location of gene variants and the clinical phenotype, a retrospective analysis was performed on 118 patients with a definitive genetic diagnosis of PFIC3. After screening by inclusion criteria, 63 cases (53.3%) were male and 55 cases (46.6%) were female. Among them, 44 cases (37.2%) had pure homozygous variants and 74 cases (62.7%) had complex heterozygous variants (Fig. 3A, B). Among all patients, 104 had two *ABCB4* locus variants, one on each allele; 13 had three *ABCB4* locus variants; and one had five *ABCB4* locus variants. Overall, 118 patients accumulated a total of 252 variants. Of these, 194 (76.9%) were missense variants, which were the most common type (Fig. 3C). There were also 25 (9.9%) frameshift variants, 21 (8.3%) nonsense variants, 9 (3.5%) splice variants, and 3 (1.1%) synonymous variants (Fig. 3C). The highest frequency of variant was in exon_7 (30 times), followed by exon_10 (22 times), exon_22 (19 times), exon_14 (18 times) and exon_15 (15 times) (Fig. 3E).

The age of onset of the patients was stratified into four stages: Infantile ($0 \leq$ age < 5 years), Childhood ($5 \leq$ age < 12 years), Adolescent ($12 \leq$ age < 18 years) and Adult (age ≥ 18 years). The results showed that the mean age of onset of the disease was 6.5 years, and 69 (58.4%) of the patients had onset in the Infantile stage, accompanied by rapid disease progression (Fig. 3D). Further analysis revealed that gender had no significant impact on the patients' age of onset.(Fig. 3F). The mean age of onset was 2.67 years for patients with pure heterozygous variants and 8.78 years for patients with complex heterozygous variants. The results of the t-test indicated a significantly earlier age of onset in patients with pure heterozygous variants compared to those with complex heterozygous variants ($p < 0.0001$), suggesting a greater pathogenicity in the pure heterozygous variants (Fig. 3G).

Random forest modeling analysis of variant sites and age of onset and clinical significance

To further elucidate the relationship between variant sites and the early or late clinical disease onset, Random Forest classification was used to analyze the correlation between the exons or introns containing variant sites and age at onset of disease, and to construct a prediction model (Fig. 3H-J). The model achieved an accuracy of 0.841 on the training set, suggesting its robust performance on known data. Meanwhile, the model achieved an accuracy of 0.667 on the test set, indicating its satisfactory generalization capabilities. The recall and precision on the test set are 0.667 and 0.726, respectively, showing that the model has a slight advantage in positive case recognition. The F1 score of 0.667 is consistent with the accuracy and recall, validating the model's overall performance on the test set. (Fig. 3H-J)

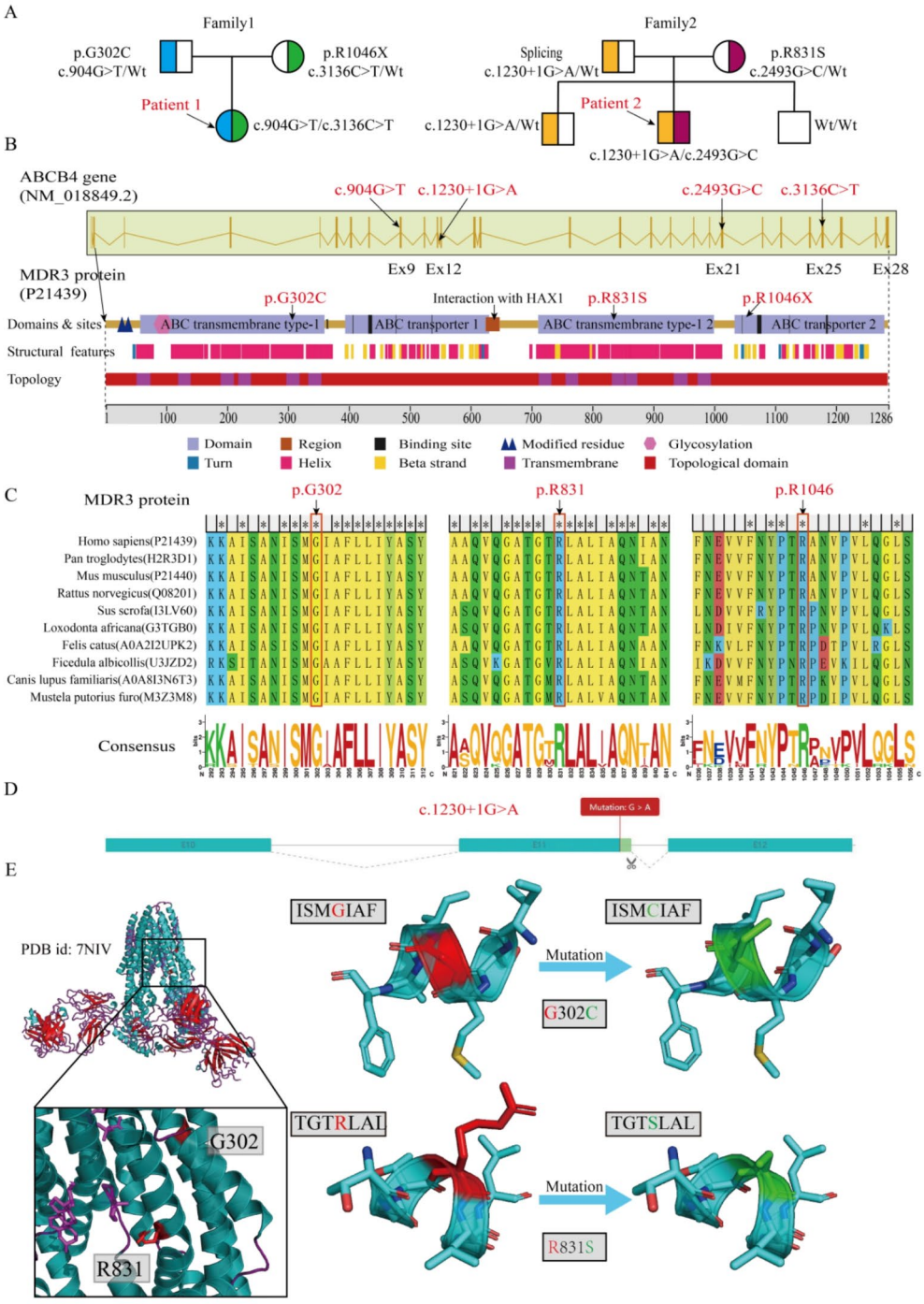


Fig. 1. Genetic Variation Analysis of the ABCB4 Gene. **(A)** Pedigrees of two patient families. Squares and circles represent male and female members, respectively. Different colors denote heterozygous carriers. Whole-genome sequencing and sequencing were performed on all family members, and genotypes are marked on the chart. **(B)** Sequencing reveals four heterozygous mutations in the ABCB4 gene: C.904G > T (p.G302C), C.1230 + 1G > A (splicing), C.2493G > C (p.R831S), and C.3136G > T (p.R1046X). **(C)** Homology alignment of the coding sequence of the MDR3 protein. All sequences are retrieved from the Ensembl database. **(D)** Prediction of the impact of the c.1230 + 1G > A variant on protein splicing. **(E)** Demonstration of tertiary structural changes in two VUS proteins.

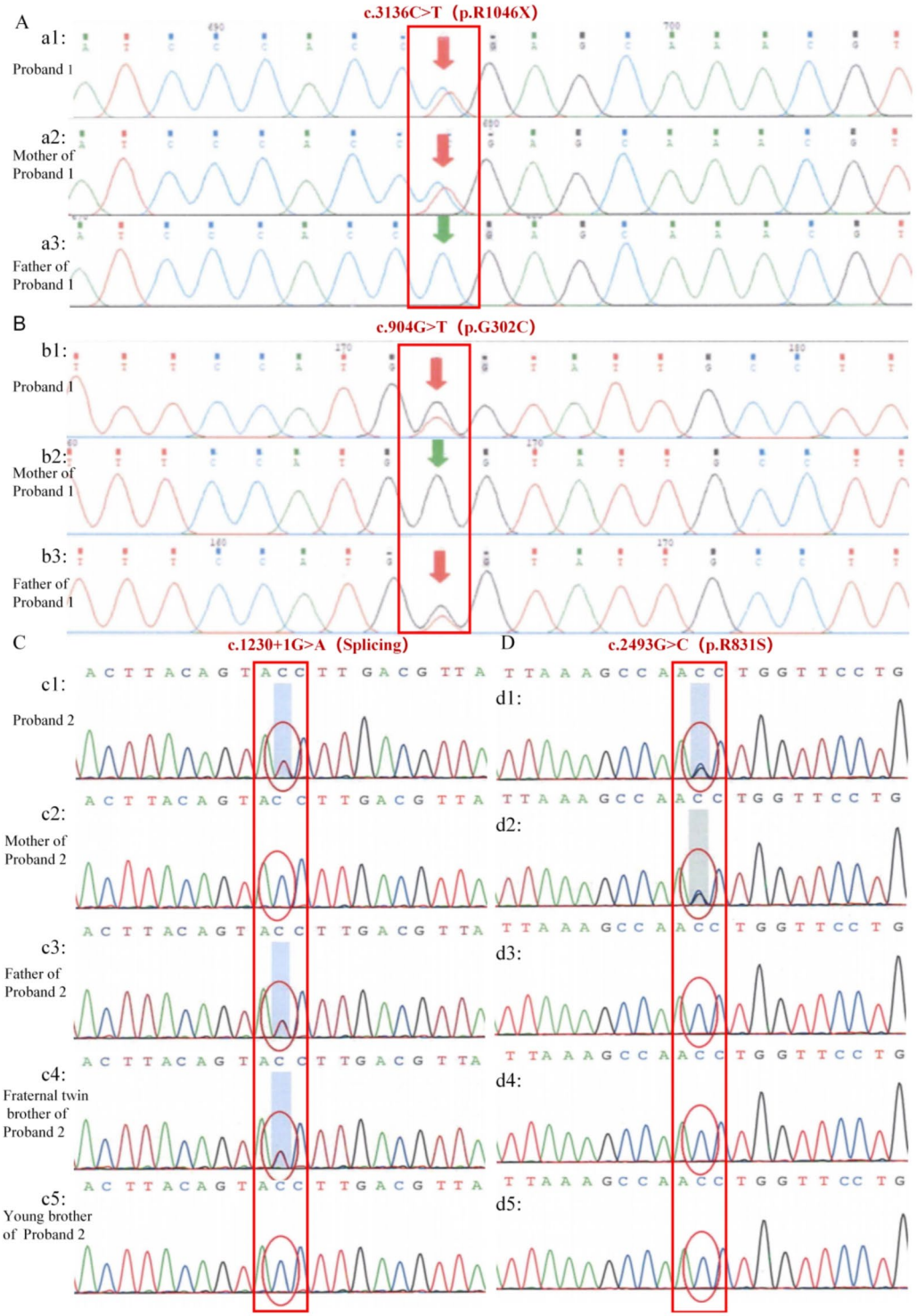
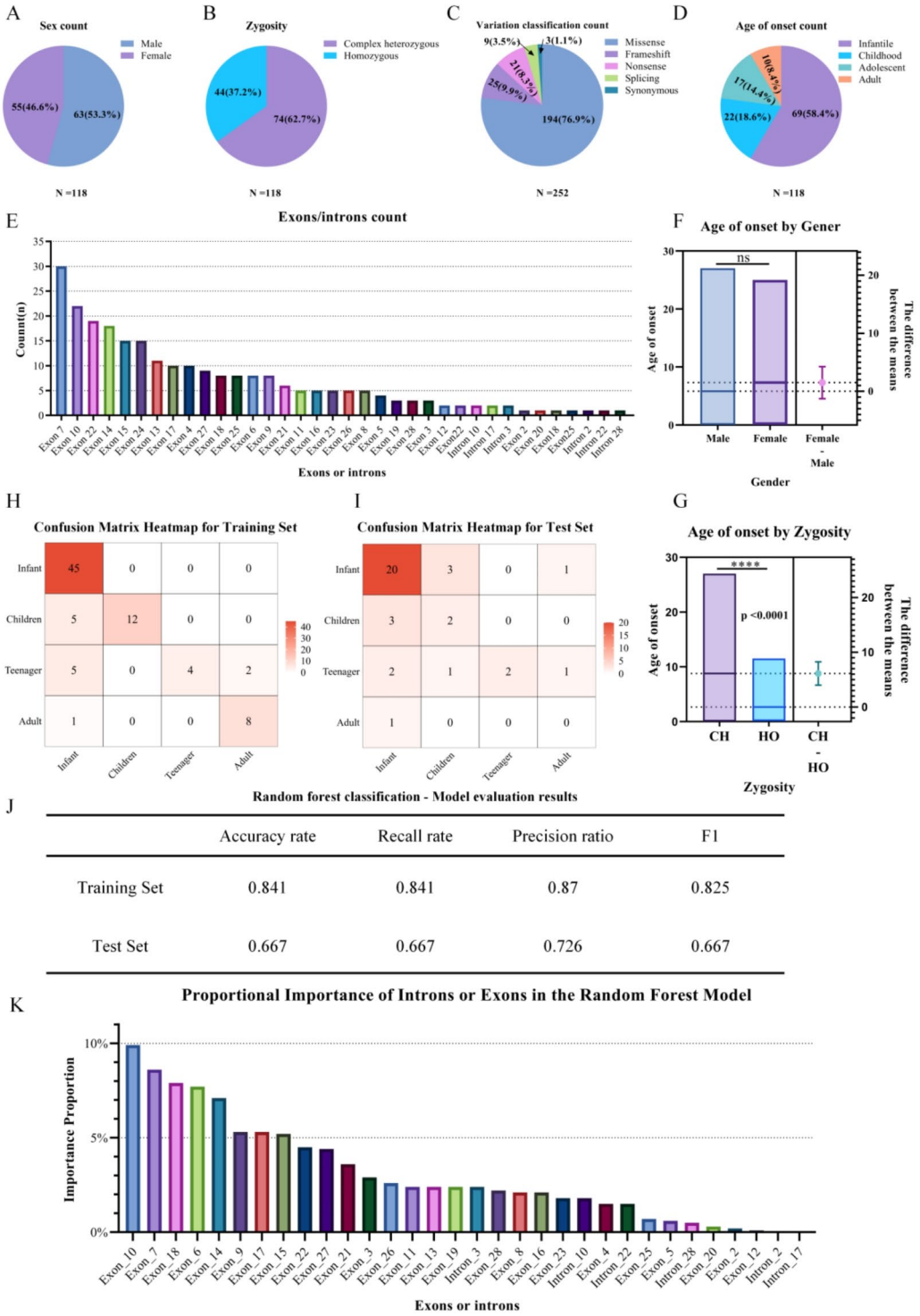


Fig. 2. Sequencing of the ABCB4 Gene in Two Patients and Their Families. **(A)** C.3136G > T (p.R1046X) site in Patient 1 and their parents. **(B)** C.904G > T (p.G302C) site in Patient 1 and their parents. **(C)** C.1230 + 1G > A (splicing) and C.2493G > C (p.R831S) sites in Patient 2, their parents, their fraternal twin brother, and their younger brother.

The localization of the variant site, specifically in an exon or intron, was deemed a pivotal variable in the model's feature extraction. The feature importance scores revealed that exon_10 had the highest score, accounting for 9.9% of the total importance. Subsequently, exon_7 (8.6%), exon_18 (7.9%), exon_6 (7.7%), exon_14 (7.1%), and exon_9 (5.3%) exhibited a significant impact on predicting the age at onset (Fig. 3K). Notably, exon_10 and



	Accuracy rate	Recall rate	Precision ratio	F1
Training Set	0.841	0.841	0.87	0.825
Test Set	0.667	0.667	0.726	0.667

Fig. 3. Sequencing of the ABCB4 Gene in Two Patients and Their Families. **(A)** Gender distribution of patients. **(B)** zygosity analysis of patients. **(C)** Proportion of variant types in patients. **(D)** Distribution of age at onset among patients. **(E)** Comparative analysis of exon and intron variation frequencies. **(F)** Age at onset by gender. **(G)** Age at onset by zygosity. Confusion matrix heatmap of variant sites across four age intervals in the **(H)** training set and **(I)** test sets. **(J)** Performance evaluation of the random forest classification model in the training and test sets. **(K)** Random forest model analysis of the importance of variant site localization.

exon_7, which exhibited high variant frequencies (22 and 30 occurrences, respectively), also ranked highly in feature importance scores, emphasizing their clinical relevance in predictive modeling (Fig. 3E, K).

These findings indicate that the localization of specific exons exerts a notable influence on the age at onset of disease among patients, particularly exon_10 and exon_7. Pure heterozygous variants are correlated with an earlier age at onset, suggesting that early detection and targeted intervention focused on these variant sites may be crucial for disease management and optimizing prognosis in clinical practice. By pinpointing these crucial variant sites, clinicians can more accurately predict the age of disease onset in patients, thereby facilitating the development of personalized treatment strategies to improve therapeutic outcomes.

Clinical management of PFIC3 patients and analysis of liver transplantation treatment effects

Considering their medical history, test results, and gene variants, both patients were definitively diagnosed with PFIC3. Despite ongoing therapy with ursodeoxycholic acid and glutathione following diagnosis, the progression of cirrhosis remained unabated. In an effort to provide relief and prolong life, patient 1 underwent orthotopic liver transplantation for end-stage liver disease at age 2, while patient 2 underwent the same procedure at age 23. On post-transplant day 16, patient 1's levels of TBil, DBil, TBA, ALT, AST, ALP and GGT normalized (Fig. 4A). During the follow-up period, patient 1's liver function and biochemical markers remained within normal ranges, with no observed abnormalities in growth and development. Patient 2 exhibited a substantial reduction in TBil, DBil, TBA, ALT and AST levels upon hospital discharge on post-transplant day 12 (Fig. 4B). Long-term follow-up revealed that all liver function tests returned to the normal range. Following liver transplantation, all indices for both patients normalized, demonstrating a remarkable therapeutic response. Retrospective analysis concurred with previous research, confirming that liver transplantation holds the potential to eradicate PFIC3.

Discussion

In this study, two PFIC3 patients from our institution and 118 PFIC3 patients reported in the literature were analyzed in detail and systematically studied. We identified that both patients carried compound heterozygous variants in the *ABCB4* gene, including three novel variants and one known variant. Multi-sequence alignment and bioinformatics prediction tools indicated that these variants likely impact the structure and function of the MDR3 protein, leading to significant clinical manifestations of liver dysfunction and the development of cholestasis. Meanwhile, liver transplantation demonstrated substantial improvement in the prognosis of both patients, reaffirming its essential role in the management and treatment of PFIC3. Random forest model analysis of the 118 patients in the literature revealed that the specific location of exons, particularly exon_10 and exon_7, significantly correlates with the age of disease onset.This finding provides new insights for early intervention and disease course prediction in PFIC3.

PFIC3 is a rare autosomal recessive genetic disease with a very low incidence. The cardinal clinical manifestations are jaundice, pruritus and abnormal liver function. As the disease progresses, patients may present with hepatosplenomegaly, cirrhosis, portal hypertension, and its complications, such as ascites and esophageal variceal bleeding[22,23]. Ursodeoxycholic acid (UDCA) has been shown to improves liver function and alleviates clinical symptoms in a subset of PFIC3 patients[24]. However, patients with advanced disease stages often depend on liver transplantation as a life-saving intervention. The prognosis of patients after liver transplantation is significantly improved, with 75–100% of patients being able to regain normal liver function and quality of life after the procedure at short-term follow-up of 3–5 years[25].

The primary cause of PFIC3 is the dysfunction of the MDR3 protein, encoded by the *ABCB4*gene[5]. MDR3 is mainly expressed on the surface of the capillary bile duct membrane of hepatocytes, and its main function is to transport phospholipids (especially phosphatidylcholine) from hepatocytes into bile[26,27]. Phospholipids form mixed micelles in the bile, and, along with bile salts and cholesterol, help to dissolve and excrete cholesterol and other fat-soluble substances[28,29]. When MDR3 is absent, non-phosphatidylcholine bile salts have a toxic detergent effect directly on capillary bile membranes, leading to cholangiocyte damage and subsequent biliary injury, biliary stasis, small ductular proliferation, inflammation, and biliary hyperplasia. These conditions can be followed by inflammatory infiltration, fibrosis in the portal area, and eventually cirrhosis and liver failure.

Patient 1, presenting at 16 months with jaundice and abdominal distension, was found to have two heterozygous variants in the *ABCB4*gene: c.3136 C > T (p.R1046X) and c.904G > T (p.G302C). The maternal variant c.3136 C > T (p.R1046X) has been reported to be associated with PFIC3 and was classified as pathogenic (PVS1 + PS4 + PM2_Supporting).This variant may cause protein truncation or nonsense-mediated mRNA degradation, affecting MDR3 function[30,31].The paternal variant c.904G > T (p.G302C) is of unknown significance and has not been reported in the literature, but is predicted to be harmful by multiple bioinformatic tools (PM3 + PM2_Supporting + PP3).The study concluded that nonsense and deletion variants, PFIC3, are more likely to have severe clinical symptoms[26]. The severe early onset in Patient 1 suggests significant impairment of MDR3 function, likely due to reduced expression from the c.3136 C > T variant and structural changes from the c.904G > T variant.

Patient 2, diagnosed with jaundice at 13, showed abnormal liver function. Despite treatment with ursodeoxycholic acid and glutathione, mild liver function abnormalities persisted. Genetic testing at age 18 revealed complex heterozygous variants in *ABCB4*: c.1230 + 1G > A (splicing) and c.2493G > C (p.R831S). These variants are novel, with c.1230 + 1G > A classified as likely pathogenic(PVS1 + PM2_Supporing) and c.2493G > C as a variant of unknown significance(PVS1 + PM2_Supporing). The adolescent onset suggests the importance of MDR3 function in early life. Bioinformatic analysis predicted significant structural changes in MDR3 due to the c.2493G > C variant. The c.1230 + 1G > A variant likely causes a frameshift and premature termination, reducing MDR3 expression.

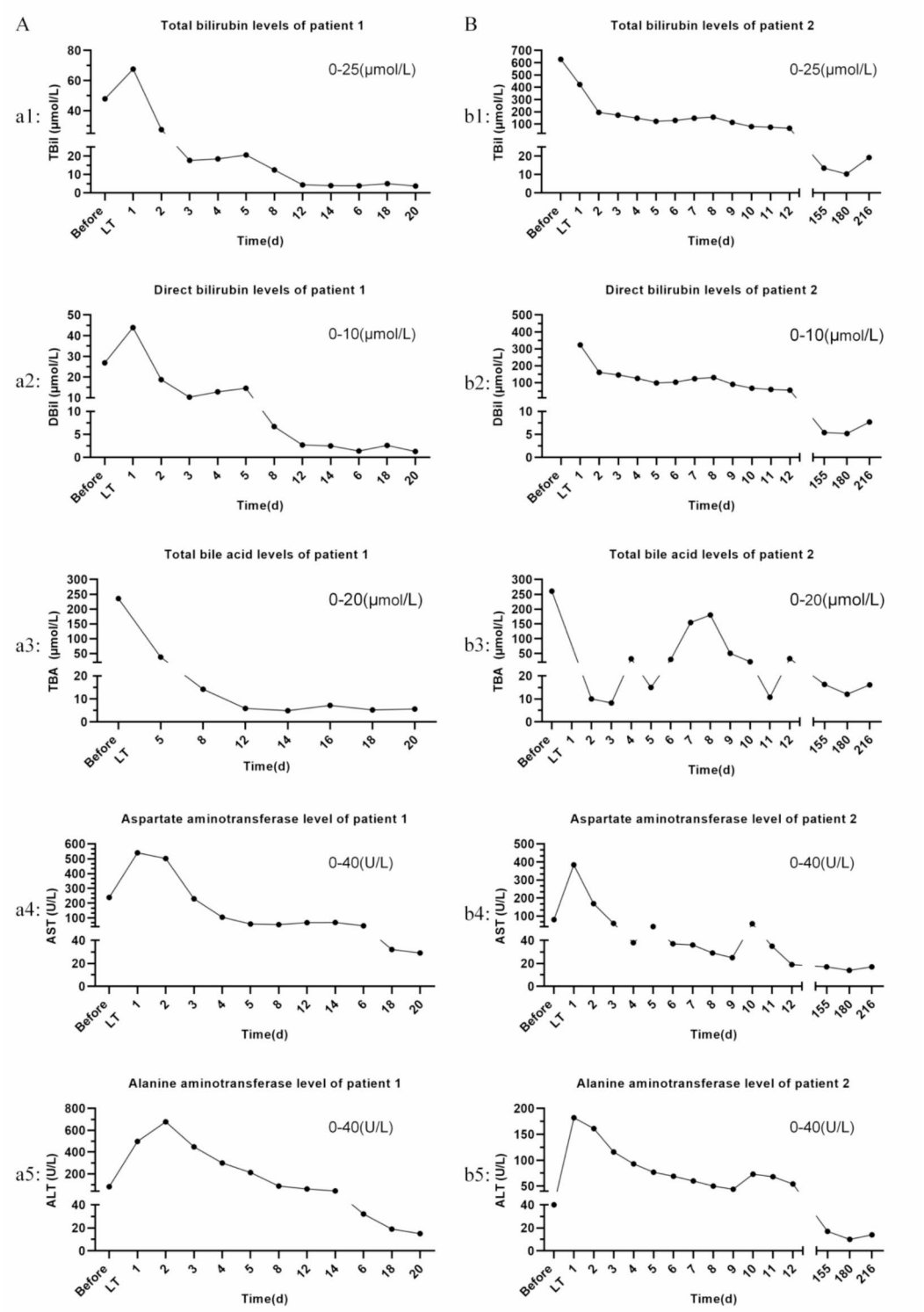


Fig. 4. Biochemical Marker Changes Before and After Liver Transplantation in PFIC3 Patients. Changes in TBil, DBil, TBA, AST, and ALT levels before and after liver transplantation in **(A)** Patient 1and **(B)** Patient 2.

Due to the limited sample size of two cases from our center, we expanded our study by incorporating data from 118 PFIC3 patients from literature. This comprehensive analysis revealed that missense variants are the most common, accounting for 76.9% of cases, with the highest frequency of variants in exons 7 and 10. The mean age of onset in patients with pure heterozygous variants is significantly earlier than in those with complex heterozygous variants, at 2.67 versus 8.78 years, respectively ($P < 0.0001$), indicating stronger pathogenicity of pure heterozygous variants. Random forest analysis confirmed that exons 10 and 7 are significant predictors of

age of onset, with the model achieving an accuracy of 0.841 on the training set and 0.667 on the test set. These findings highlight the importance of targeting specific exon locations for early detection and intervention in clinical practice, improving disease management and prognosis. Identifying key variant sites enables clinicians to predict the age of onset and develop personalized treatment plans, enhancing therapeutic outcomes. This study provides novel insights into the early diagnosis and personalized treatment of PFIC3, emphasizing the importance of early detection and intervention, and laying a foundation for future clinical management strategies.

Diagnosis of PFIC3 primarily involves genetic analysis to detect *ABCB4*gene variants. Blood tests often show impaired hepatic function with elevated bile acid levels[32].Imaging techniques like ultrasonography and magnetic resonance cholangiopancreatography (MRCP) assess the liver and biliary tract. Liver biopsy can reveal bile duct injury and fibrosis[33]. Advances in genomics, molecular biology, and artificial intelligence may lead to intelligent diagnostic systems, enhancing diagnostic accuracy and efficiency.

PFIC3 treatment strategies include medications like UDCA to improve bile flow and reduce symptoms[34]. UDCA reduces bile duct epithelial cell damage by increasing water-soluble bile components, improving symptoms and liver function[35]. For severe cholestasis, partial external biliary drainage (PEBD) can reduce bile duct pressure and injury[36]. Liver transplantation is the only effective treatment for end-stage liver disease[37]. Both patients in this study showed rapid liver function normalization post-transplant, with Patient 1 experiencing normal growth and development, demonstrating the efficacy of liver transplantation in correcting metabolic defects in *ABCB4* variant patients.

Advancements in gene therapy and novel therapeutic approaches are promising for PFIC3 treatment. These new therapies are expected to provide more effective options for patients[12,38]. Additionally, The development of personalized treatment strategies, based on individual genetic variations and clinical manifestations, is a crucial research direction. This approach will enhance treatment efficacy[38,39]. Despite challenges, technological progress in PFIC3 research offers hope for more effective and diverse treatment options, helping patients overcome the disease and regain health.

Conclusion

This study underscores the critical role of genetic analysis in diagnosing PFIC3 and highlights the significant impact of specific *ABCB4* gene variants on disease onset and progression. The analysis of both institutional cases and literature data reveals valuable insights into the genetic underpinnings of PFIC3, emphasizing the importance of early detection and personalized treatment strategies. Liver transplantation remains a cornerstone in managing advanced PFIC3, significantly improving patient outcomes. Future research should focus on advancing gene therapy and developing intelligent diagnostic systems to enhance treatment efficacy and provide patients with diverse therapeutic options.

Data availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request. Researchers interested in accessing the data should contact Xiaofang Zhao at Email: zhaoxiaofang2099@163.com .

Received: 7 August 2024; Accepted: 6 November 2024
Published online: 09 November 2024

References

1. Jacquemin, E. et al. The wide spectrum of multidrug resistance 3 deficiency: from neonatal cholestasis to cirrhosis of adulthood. *Gastroenterology*. **120**, 1448–1458. https://doi.org/10.1053/gast.2001.23984 (2001).
2. Reichert, M. C. & Lammert, F. *ABCB4* gene aberrations in human liver disanse: An evolving spectrum. *Semin Liver Dis.* **38**, 299–307. https://doi.org/10.1055/s-0038-1667299 (2018).
3. Davit-Spraul, A. et al. The spectrum of liver diseases related to *ABCB4* gene mutations: pathophysiology and clinical aspects. *Semin Liver Dis.* **30**, 134–146. https://doi.org/10.1055/s-0030-1253223 (2010).
4. Vitale, G. et al. Familial intrahepatic cholestasis: New and wide perspectives. *Dig. Liver Dis.* **51**, 922–933. https://doi.org/10.1016/j.dld.2019.04.013 (2019).
5. de Vree, J. M. et al. Mutations in the *MDR3* gene cause progressive familial intrahepatic cholestasis. *Proc. Natl. Acad. Sci. USA*. **95**, 282–287. https://doi.org/10.1073/pnas.95.1.282 (1998).
6. Smit, J. J. M., Schinkel, A. H. & Elferink, O. Homozygous disruption of the murine *mdr2* P-Glycoprotein gene leads to a complete absence of phospholipid from bile and to liver disease. *Cell*. **75**, 451–462. https://doi.org/10.1016/0092-8674(93)90380-9 (1993).
7. Ruetz, S. & Gros, P. Phosphatidylcholine translocase: a physiological role for the *mdr2* gene. *Cell*. **77**, 1071–1081. https://doi.org/10.1016/0092-8674(94)90446-4 (1994).
8. Oude Elferink, R. P. J., Paulusma, C. C. & Groen, A. K. Hepatocanalicular transport defects: pathophysiologic mechanisms of rare diseases. *Gastroenterology*. **130**, 908–925. https://doi.org/10.1053/j.gastro.2005.08.052 (2006).
9. Hirschfield, G. M. et al. The genetics of complex cholestatic disorders. *Gastroenterology*. **144**, 1357–1374. https://doi.org/10.1053/j.gastro.2013.03.053 (2013).
10. Beuers, U. et al. New paradigms in the treatment of hepatic cholestasis: from UDCA to FXR, PXR and beyond. *J. Hepatol.* **62**, 25–S37. https://doi.org/10.1016/j.jhep.2015.02.023 (2015).
11. Siew, S. M. et al. Prevention of cholestatic liver disease and reduced tumorigenicity in a murine model of PFIC type 3 using hybrid AAV-*piggyBac* gene therapy. *Hepatology*. **70**, 2047–2061. https://doi.org/10.1002/hep.30773 (2019).
12. Wei, G. et al. Synthetic human *ABCB4* mRNA therapy rescues severe liver disease phenotype in a BALB/c.*Abcb4*$^{-/-}$ mouse model of PFIC3. *J. Hepatol.* **74**, 1416–1428. https://doi.org/10.1016/j.jhep.2020.12.010 (2021).
13. Englert, C. et al. Liver transplantation in children with progressive familial intrahepatic cholestasis. *Transplantation*. **84**, 1361–1363. https://doi.org/10.1097/01.tp.0000282869.94152.4f (2007).
14. Saleem, K. et al. Evaluation of a novel missense mutation in *ABCB4* gene causing progressive familial intrahepatic cholestasis type 3. *Dis Markers* 6292818 (2020). (2020). https://doi.org/10.1155/2020/6292818

15. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424. https://doi.org/10.1038/gim.2015.30 (2015).
16. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081. https://doi.org/10.1038/nprot.2009.86 (2009).
17. Adzhubei, I., M Jordan, D. & R Sunyaev, S. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet. Chap.* **7**, Unit720. https://doi.org/10.1002/0471142905.hg0720s76 (2013).
18. Schwarz, J. M. et al. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods.* **7**, 575–576. https://doi.org/10.1038/nmeth0810-575 (2010).
19. Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* **31**, 2745–2747. https://doi.org/10.1093/bioinformatics/btv195 (2015).
20. O. Desmet, F. et al. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37** (e67). https://doi.org/10.1093/nar/gkp215 (2009).
21. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027. https://doi.org/10.1093/molbev/msab120 (2021).
22. Bull, L. N. & Thompson, R. J. Progressive familial intrahepatic cholestasis. *Clin. Liver Dis.* **22**, 657–669. https://doi.org/10.1016/j.cld.2018.06.003 (2018).
23. Oude Elferink, R. P. J. & Paulusma, C. C. Function and pathophysiological importance of ABCB4 (MDR3 P-glycoprotein). *Pflugers Arch.* **453**, 601–610. https://doi.org/10.1007/s00424-006-0062-9 (2007).
24. Jacquemin, E. et al. Ursodeoxycholic acid therapy in pediatric patients with progressive familial intrahepatic cholestasis. *Hepatology.* **25**, 519–523. https://doi.org/10.1002/hep.510250303 (1997).
25. Hori, T. et al. Progressive familial intrahepatic cholestasis: a single-center experience of living-donor liver transplantation during two decades in Japan. *Clin. Transpl.* **25**, 776–785. https://doi.org/10.1111/j.1399-0012.2010.01368.x (2011).
26. Keitel, V. et al. Expression and localization of hepatobiliary transport proteins in progressive familial intrahepatic cholestasis. *Hepatology.* **41**, 1160–1172. https://doi.org/10.1002/hep.20682 (2005).
27. Zolnerciks, J. K. et al. Structure of ABC transporters. *Essays Biochem.* **50**, 43–61. https://doi.org/10.1042/BSE0500043 (2011).
28. Hofmann, A. F. Bile acid secretion, bile flow and biliary lipid secretion in humans. *Hepatology.* **12**, 17S–22S (1990). discussion 22S-25S.
29. Srivastava, A. Progressive familial intrahepatic cholestasis. *J. Clin. Exp. Hepatol.* **4**, 25–36. https://doi.org/10.1016/j.jceh.2013.10.005 (2014).
30. Nakken, K. E. et al. *ABCB4* sequence variations in young adults with cholesterol gallstone disease. *Liver Int.* **29**, 743–747. https://doi.org/10.1111/j.1478-3231.2008.01914.x (2009).
31. de Vries, E. et al. Carriers of *ABCB4* gene variants show a mild clinical course, but impaired quality of life and limited risk for cholangiocarcinoma. *Liver Int.* **40**, 3042–3050. https://doi.org/10.1111/LIV.14662 (2020).
32. Gordo-Gilart, R. et al. Heterozygous *ABCB4* mutations in children with cholestatic liver disease. *Liver Int.* **36**, 258–267. https://doi.org/10.1111/liv.12910 (2016).
33. Poupon, R. et al. Combined features of low phospholipid-associated cholelithiasis and progressive familial intrahepatic cholestasis 3. *Liver Int.* **30**, 327–331. https://doi.org/10.1111/j.1478-3231.2009.02148.x (2009).
34. Delaunay, J. L. et al. Functional defect of variants in the ATP-binding sites of ABCB4 and their rescue by the CFTR potentiator, ivacaftor (VX-770). *Hepatology.* **65**, 560–570. https://doi.org/10.1002/hep.28929 (2017).
35. Gordo-Gilart, R. et al. Functional analysis of *ABCB4* mutations relates clinical outcomes of progressive familial intrahepatic cholestasis type 3 to the degree of MDR3 floppase activity. *Gut.* **64**, 147–155. https://doi.org/10.1136/gutjnl-2014-306896 (2015).
36. Stapelbroek, J. M. et al. Liver disease associated with canalicular transport defects: current and future therapies. *J. Hepatol.* **52**, 258–271. https://doi.org/10.1016/j.jhep.2009.11.012 (2010).
37. Liu, C. et al. Novel resequencing chip customized to diagnose mutations in patients with inherited syndromes of intrahepatic cholestasis. *Gastroenterology.* **132**, 119–126. https://doi.org/10.1053/j.gastro.2006.10.034 (2007).
38. Aronson, S. J. et al. Liver-directed gene therapy results in long-term correction of progressive familial intrahepatic cholestasis type 3 in mice. *J. Hepatol.* **71**, 153–162. https://doi.org/10.1016/j.jhep.2019.03.021 (2019).
39. Weber, N. D., Martínez-García, J. & González-Aseguinolaza, G. Comment on Synthetic human ABCB4 mRNA therapy rescues severe liver disease phenotype in a BALB/c.Abcb4$^{-/-}$ mouse model of PFIC3. *J. Hepatol.* **76**, 749–751. https://doi.org/10.1016/j.jhep.2021.09.033 (2022).

Acknowledgements

We thank our research participants and their families.

Author contributions

S.Y.W. and L.L.D. designed the study. S.Y.W. and Q.L. collected and analyzed clinical and genetic data. S.Y.W. and X.Y.S used computer software to analyze the variation. S.Y.W.and W.W.J wrote the manuscript. L.L.D. and X.F.Z. made substantial contribution during study conception and manuscript improvements.All authors read and approved the final manuscript.

Funding

This study was supported by the Natural Science Foundation of Henan Province (242300420070); the Open Project of Key Laboratory of the Ministry of Education (2022-MEKLLC-ZD-002); the National Natural Science Foundation of China (81902904).

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (Approval No. 2022-KY-0826-001), and all family members or legal guardians of patients under 18 years of age signed a written informed consent. The data in this study were anonymized, or treated confidentially, and did not affect the diagnosis or treatment of the patients.

Consent for publication

Not applicable.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-79123-6.

Correspondence and requests for materials should be addressed to L.D. or X.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.